

On the Error Bound in a Combinatorial Central Limit Theorem

Louis H.Y. Chen¹ and Xiao Fang²

¹*Department of Mathematics
National University of Singapore
10 Lower Kent Ridge Road
Singapore 119076
Republic of Singapore
e-mail: matchyl@nus.edu.sg*

²*Department of Statistics and Applied Probability
National University of Singapore
6 Science Drive 2
Singapore 117546
Republic of Singapore
e-mail: fangwd2003@gmail.com*

Abstract: Let $\mathbb{X} = \{X_{ij} : 1 \leq i, j \leq n\}$ be an $n \times n$ array of independent random variables. Let π be a uniform random permutation of $\{1, 2, \dots, n\}$, independent of \mathbb{X} , and let $W = \sum_{i=1}^n X_{i\pi(i)}$. Suppose \mathbb{X} is standardized so that $\mathbb{E}W = 0$, $\text{Var}(W) = 1$. We prove that the Kolmogorov distance between the distribution of W and the standard normal distribution is bounded by $447 \sum_{i,j=1}^n \mathbb{E}|X_{ij}|^3/n$. Our approach is by Stein's method of exchangeable pairs and the use of a concentration inequality.

AMS 2000 subject classifications: 60F05.

Keywords and phrases: combinatorial central limit theorem, concentration inequality, exchangeable pairs, Stein's method.

1. Introduction and statement of the main result

Motivated by permutation tests in non-parametric statistics, Wald and Wolfowitz [20] proved a central limit theorem for the combinatorial statistics $\sum_{i=1}^n a_i b_{\pi(i)}$ where $\{a_i, b_j : i, j \in [n] := \{1, 2, \dots, n\}\}$ are real numbers and π is a uniform random permutation of $[n]$. Their result was generalized to real arrays $\{c_{ij} : i, j \in [n]\}$ by Hoeffding [14]. Extension to random

arrays $\{X_{ij} : i, j \in [n]\}$ where (X_{ij}) are independent random variables was considered by Ho and Chen [13]. Using the concentration inequality approach in Stein's method, they proved a bound on the Kolmogorov distance between the distribution of $\sum_{i=1}^n X_{i\pi(i)}$ and the normal distribution with the same mean and variance. The bound in [13] is optimal only when $|X_{ij}| \leq C$ for an absolute constant C . A third-moment bound for a combinatorial central limit theorem for real arrays $\{c_{ij} : i, j \in [n]\}$ was obtained by Bolthausen [2], who used Stein's method and induction. However, the absolute constant in the bound in [2] is not explicit. A bound with an explicit constant for real arrays with $|c_{ij}| \leq C$ was obtained by Goldstein [11] using Stein's method and zero-bias coupling (see also [6]). Under the same setting as Ho and Chen [13], Neammanee and Suntornchost [16] stated a third-moment bound. They used the same Stein identity in [13], which dates back to Chen [3], and the concentration inequality approach. However, there is an error in the proof in [16], where the first equality and the second inequality on page 576 are incorrect because of the dependence among $S(\tau)$, ΔS and $M(t)$.

In this paper, we give a different proof of the combinatorial central limit theorem. Our result gives a third-moment bound with an explicit constant under the setting of Ho and Chen. Our approach is by Stein's method of exchangeable pairs and the use of a concentration inequality. The use of an exchangeable pair simplifies the construction of a Stein identity as compared to the construction in [13] and [16].

Stein's method was introduced by Stein ([18], [19]) and has become a popular tool in proving distributional approximation results because of its power in handling dependence among random variables. We refer to Barbour and Chen [1] and Chen, Goldstein and Shao [6] for an introduction of Stein's method. The notion of exchangeable pair was introduced by Stein [19], and first used in Diaconis [10]. The concentration inequality approach was also introduced by Stein (see [13]) and was developed by Chen ([4], [5]) and Chen and Shao ([7], [8]). This approach provides a smoothing technique and a way of obtaining third-moment bounds on the Kolmogorov distance.

The following is our main result.

Theorem 1.1. *Let $\mathbb{X} = \{X_{ij} : i, j \in [n]\}$ be an $n \times n$ array of independent random variables*

where $n \geq 1$, $\mathbb{E}X_{ij} = c_{ij}$, $\text{Var}X_{ij} = \sigma_{ij}^2$. Assume

$$c_{i.} = c_{.j} = c_{..} = 0 \quad (1.1)$$

where $c_{i.} = \sum_{j=1}^n c_{ij}/n$, $c_{.j} = \sum_{i=1}^n c_{ij}/n$, $c_{..} = \sum_{i,j=1}^n c_{ij}/n^2$. Let π be a uniform random permutation of $[n]$, independent of \mathbb{X} , and let $W = \sum_{i=1}^n X_{i\pi(i)}$. Then

$$(1) \quad \text{Var}(W) = \frac{1}{n} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{n-1} \sum_{i,j=1}^n c_{ij}^2, \quad (1.2)$$

(2) assuming $\text{Var}(W) = 1$, we have

$$\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \leq 447\gamma \quad (1.3)$$

where Φ is the standard normal distribution function and

$$\gamma = \frac{1}{n} \sum_{i,j=1}^n \mathbb{E}|X_{ij}|^3. \quad (1.4)$$

For real arrays, we have the following corollary of Theorem 1.1 by letting $\sigma_{ij} = 0$.

Corollary 1.2. Let $\{c_{ij} : i, j \in [n]\}$ be an $n \times n$ real array. Assume

$$c_{i.} = c_{.j} = c_{..} = 0$$

where $c_{i.} = \sum_{j=1}^n c_{ij}/n$, $c_{.j} = \sum_{i=1}^n c_{ij}/n$, $c_{..} = \sum_{i,j=1}^n c_{ij}/n^2$, and assume

$$\frac{1}{n-1} \sum_{i,j=1}^n c_{ij}^2 = 1.$$

Let π be a uniform random permutation of $[n]$, and let $W = \sum_{i=1}^n c_{i\pi(i)}$. Then we have

$$\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \frac{447}{n} \sum_{i,j=1}^n |c_{ij}|^3. \quad (1.5)$$

Moreover, if $|c_{ij}| \leq C$ for all $i, j \in [n]$, then

$$\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \leq 447C. \quad (1.6)$$

We remark that the constant 447 in Corollary 1.2 can be reduced by a modification of the proof of Theorem 1.1 for this special case. The error bound in (1.5) was obtained by

Bolthausen [2] except that the constant in his bound is not explicit. The error bound in (1.6) is of the same order as that in [11] and in Theorem 6.1 of [6], although the constant in Theorem 6.1 is smaller than that in (1.6).

For any $n \times n$ array of independent random variables, we can standardize it and apply Theorem 1.1 to obtain the following corollary.

Corollary 1.3. *Let $\mathbb{Y} = \{Y_{ij} : i, j \in [n]\}$ be an $n \times n$ array of independent random variables where $n \geq 1$, $\mathbb{E}Y_{ij} = \mu_{ij}$, $\text{Var}Y_{ij} = \sigma_{ij}^2$. Let π be a uniform random permutation of $[n]$, independent of \mathbb{Y} , and let $V = \sum_{i=1}^n Y_{i\pi(i)}$. Then we have*

$$\sup_{z \in \mathbb{R}} \left| P\left(\frac{V - n\mu_{..}}{\sigma} \leq z\right) - \Phi(z) \right| \leq \frac{447}{n} \sum_{i,j=1}^n \mathbb{E}|Y_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}|^3 / \sigma^3$$

where

$$\begin{aligned} \mu_{i.} &= \sum_{j=1}^n \mu_{ij} / n, & \mu_{.j} &= \sum_{i=1}^n \mu_{ij} / n, & \mu_{..} &= \sum_{i,j=1}^n \mu_{ij} / n^2, \\ \sigma^2 &= \text{Var}(V) = \frac{1}{n} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{n-1} \sum_{i,j=1}^n (\mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..})^2, \end{aligned}$$

and Φ is the standard normal distribution function.

In [21], as a step in the algorithm which generates a random mapping in the Johnson-Lindenstrauss lemma [15], the following problem is considered. Let $\{Y_1, \dots, Y_n\}$ be independent mean zero random variables. For a positive integer $k \leq n$, let $\{Y_{\xi_1}, \dots, Y_{\xi_k}\}$ be uniformly chosen from $\{Y_1, \dots, Y_n\}$ without replacement, and let $V = \sum_{i=1}^k Y_{\xi_i}$. Suppose $\{Y_1, \dots, Y_n\}$ are standardized so that $\sum_{i=1}^n \mathbb{E}Y_i^2 = n/k$, which is equivalent to $\text{Var}(V) = 1$. In [21], it is shown that $|V|$ is well concentrated around $\mathbb{E}|V|$, therefore an estimate of $\mathbb{E}|V|$ is of interest. Using Stein's method, the bound

$$|\mathbb{E}|V| - \sqrt{\frac{2}{\pi}}| \leq \frac{3k \sum_{i=1}^n \mathbb{E}|Y_i|^3}{2n}$$

is proved in [21] as a corollary of a general bound on the Wasserstein distance between $\mathcal{L}(V)$ and $N(0, 1)$. Using the same idea as in the Poisson approximation for the hypergeometric distribution in Corollary 3.4 of [3], we let the $n \times n$ array \mathbb{Y} be such that the first k rows are independent copies of $\{Y_1, \dots, Y_n\}$ and the other rows are zeros. Then $\mathcal{L}(\sum_{i=1}^n Y_{i\pi(i)}) = \mathcal{L}(\sum_{i=1}^k Y_{\xi_i})$. Therefore we have the following result by Corollary 1.3.

Corollary 1.4. *Let $\{Y_1, \dots, Y_n\}$ be independent random variables with $\mathbb{E}Y_i = \mu_i$, $\text{Var}(Y_i) = \sigma_i^2$. For a positive integer $k \leq n$, let $\{Y_{\xi_1}, \dots, Y_{\xi_k}\}$ be uniformly chosen from $\{Y_1, \dots, Y_n\}$ without replacement, and let $V = \sum_{i=1}^k Y_{\xi_i}$. Then we have*

$$\begin{aligned} & \sup_{z \in \mathbb{R}} |P(\frac{V - k\bar{\mu}}{\sigma} \leq z) - \Phi(z)| \\ & \leq \frac{447}{n\sigma^3} \left[k \sum_{i=1}^n \mathbb{E} \left| \frac{k}{n}(Y_i - \mu_i) + \frac{n-k}{n}(Y_i - \bar{\mu}) \right|^3 + (n-k) \sum_{i=1}^n \left| \frac{k}{n}(\mu_i - \bar{\mu}) \right|^3 \right] \end{aligned} \quad (1.7)$$

where

$$\bar{\mu} = \sum_{i=1}^n \mu_i / n, \quad \sigma^2 = \text{Var}(V) = \frac{k}{n} \sum_{i=1}^n \sigma_i^2 + \frac{k(n-k)}{n(n-1)} \sum_{i=1}^n (\mu_i - \bar{\mu})^2,$$

and Φ is the standard normal distribution function.

If $\mu_i = 0$ for all $i \in [n]$, then $\sigma^2 = \frac{k}{n} \sum_{i=1}^n \sigma_i^2$ and the error bound in (1.7) reduces to $\frac{447k}{n\sigma^3} \sum_{i=1}^n \mathbb{E}|Y_i|^3$, which is the same as the Wasserstein distance bound obtained in [21] except for the constant.

The case where $\sigma_i = 0$ for all $i \in [n]$ in Corollary 1.4 was considered by Goldstein [12] using zero-bias coupling, where a similar bound with a smaller constant was proved for the Wasserstein distance in simple random sampling (see Theorem 5.1 of [12]). Although we will not consider Wasserstein distance in this paper, we wish to mention that a Wasserstein distance bound can be obtained for the normal approximation for $\mathcal{L}(W)$ where W is defined in Theorem 1.1. The proof for the bound will not require a concentration inequality and the bound will have a smaller constant. See Section 2 of [13] for a discussion.

In the next section, we prove a concentration inequality using exchangeable pairs (Lemma 2.1) and apply it to random variables with a combinatorial structure similar to that of W in Theorem 1.1. In Section 3, we prove the main result Theorem 1.1 by Stein's method of exchangeable pairs and the concentration inequality approach.

2. Concentration inequalities via exchangeable pairs

The next lemma provides a bound on $\mathbb{P}(S \in [a, b])$ assuming the existence of an exchangeable pair (S, S') and an approximate linearity condition.

Lemma 2.1. *Suppose (S, S') is an exchangeable pair of square integrable random variables and satisfies the following approximate linearity condition*

$$\mathbb{E}(S' - S|S) = -\lambda S + R \quad (2.1)$$

for a positive number λ and a random variable R . Then, for $a < b$,

$$\begin{aligned} P(S \in [a, b]) &\leq \frac{\mathbb{E}|S| + \mathbb{E}|R|/\lambda}{\mathbb{E}S^2 - \mathbb{E}|SR|/\lambda - 1/2} \left(\frac{b-a}{2} + \delta \right) \\ &\quad + \frac{1}{\mathbb{E}S^2 - \mathbb{E}|SR|/\lambda - 1/2} \sqrt{\text{Var} \left(\mathbb{E} \left(\frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) | S \right) \right)} \end{aligned} \quad (2.2)$$

where

$$\delta = \frac{\mathbb{E}|S' - S|^3}{\lambda} \quad (2.3)$$

provided that $\mathbb{E}S^2 - \mathbb{E}|SR|/\lambda - 1/2 > 0$.

Remark 2.2. If $R = 0$ in (2.1) and $\mathbb{E}S^2 = 1$, the bound on the right-hand side of (2.2) becomes

$$b - a + 2\delta + 2\sqrt{\text{Var} \left(\mathbb{E} \left(\frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) | S \right) \right)}.$$

Bounding the last term involves studying the conditional distribution of $(S' - S)^2$ given S , which is common in the literature of Stein's method of exchangeable pairs. An advantage of our bound is that it truncates $|S' - S|$ at δ thus allowing us to keep within third moments.

Proof. Assume $\delta < \infty$ without loss of generality. From the exchangeability of S and S' ,

$$\mathbb{E}(S' - S)(f(S') + f(S)) = 0 \quad (2.4)$$

for all f such that the above expectation exists. Therefore,

$$\mathbb{E}(S' - S)(f(S') - f(S)) = 2\mathbb{E}(S - S')f(S).$$

Using the approximate linearity condition (2.1) for the right-hand side of the above equation, we have for absolutely continuous f ,

$$\mathbb{E}Sf(S) = \frac{1}{2\lambda}\mathbb{E}(S' - S) \int_0^{S'-S} f'(S+t)dt + \frac{1}{\lambda}\mathbb{E}Rf(S). \quad (2.5)$$

The identity (2.4) was introduced by Stein [19] and (2.5) was obtained by Stein [19] in the case $R = 0$ and by Rinott and Rotar [17] for $R \neq 0$.

Let f be such that $f'(w) = I(a - \delta \leq w \leq b + \delta)$ and $f(\frac{a+b}{2}) = 0$. Therefore, $|f| \leq \frac{b-a}{2} + \delta$. Using the property that for all $w, w' \in \mathbb{R}$,

$$(w' - w) \int_0^{w'-w} f'(w+t) dt \geq 0,$$

we have

$$\begin{aligned} & \frac{1}{2\lambda} \mathbb{E}(S' - S) \int_0^{S'-S} f'(S+t) dt \\ & \geq \frac{1}{2\lambda} \mathbb{E}(S' - S) \int_0^{S'-S} f'(S+t) dt I(|S' - S| \leq \delta) I(S \in [a, b]) \\ & = \mathbb{E}I(S \in [a, b]) \frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) \\ & = \mathbb{E}I(S \in [a, b]) \left[\mathbb{E} \left(\frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) | S \right) - \mathbb{E} \frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) \right] \\ & \quad + \mathbb{E}I(S \in [a, b]) \mathbb{E} \frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) \\ & := R_1 + R_2. \end{aligned}$$

Using the Cauchy-Schwarz inequality,

$$|R_1| \leq \sqrt{\text{Var} \left(\mathbb{E} \left(\frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) | S \right) \right)}.$$

From (2.1),

$$\mathbb{E}(S' - S)^2 = 2\mathbb{E}S(\lambda S - R) = 2\lambda\mathbb{E}S^2 - 2\mathbb{E}SR.$$

Therefore,

$$\begin{aligned} R_2 &= P(S \in [a, b]) \mathbb{E} \frac{1}{2\lambda} (S' - S)^2 - P(S \in [a, b]) \mathbb{E} \frac{1}{2\lambda} (S' - S)^2 I(|S' - S| > \delta) \\ &\geq P(S \in [a, b]) \left(\mathbb{E}S^2 - \frac{\mathbb{E}SR}{\lambda} \right) - P(S \in [a, b]) \frac{1}{\delta} \frac{\mathbb{E}|S' - S|^3}{2\lambda} \\ &= P(S \in [a, b]) \left(\mathbb{E}S^2 - \frac{\mathbb{E}SR}{\lambda} - \frac{1}{2} \right) \\ &\geq P(S \in [a, b]) \left(\mathbb{E}S^2 - \frac{\mathbb{E}|SR|}{\lambda} - \frac{1}{2} \right) \end{aligned}$$

where in the last equality we used the definition of δ in (2.3). Using the fact that $|f| \leq \frac{b-a}{2} + \delta$, we have

$$|\mathbb{E}Sf(S)| \leq \left(\frac{b-a}{2} + \delta \right) \mathbb{E}|S|, \quad \left| \frac{1}{\lambda} \mathbb{E}Rf(S) \right| \leq \left(\frac{b-a}{2} + \delta \right) \frac{\mathbb{E}|R|}{\lambda}.$$

The lemma is proved by applying all the above bounds to (2.5). \square

Now we apply Lemma 2.1 to establish a concentration inequality for a sum S which is defined as follows. Let \mathbb{X} be the $n \times n$ array defined in Theorem 1.1 satisfying (1.1) and

$$\frac{1}{n} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{n-1} \sum_{i,j=1}^n c_{ij}^2 = 1. \quad (2.6)$$

For $n \geq 6$. Let $\mathcal{I}, \mathcal{J} \subset [n]$ with cardinalities $|\mathcal{I}| = |\mathcal{J}| = m \in \{2, 3, 4\}$. We remove rows \mathcal{I} and columns \mathcal{J} from the original array \mathbb{X} . Since the argument in the proof will not depend on the choices of \mathcal{I} and \mathcal{J} given m , we assume the rows and columns removed are the last m rows and m columns for simplification of notation. Let τ be an independent uniform random permutation of $[n-m]$. Define the variable S by

$$S = \sum_{i=1}^{n-m} X_{i\tau(i)}. \quad (2.7)$$

Proposition 2.3. *Let S be defined by (2.7) for some $m \in \{2, 3, 4\}$. Suppose $\gamma \leq 1/c_0$ where γ was defined in (1.4), and c_0 and $n \geq 6$ are large enough to satisfy*

$$\theta := \frac{1}{2} - \frac{2n}{(n-4)c_0^{2/3}} - \frac{24n}{(n-5)^2} - \frac{2\sqrt{n}}{n-4} \sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2}} > 0. \quad (2.8)$$

Then for all $a < b$,

$$\mathbb{P}(S \in [a, b]) \leq c_1(b-a) + c_2\gamma \quad (2.9)$$

where

$$c_1 = \left(\frac{1}{2} \sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2}} + \frac{\sqrt{n}}{n-4} \right) / \theta \quad (2.10)$$

and

$$c_2 = \frac{64n}{n-4} c_1 + \sqrt{\left[\frac{8n}{(n-4)^2} + \frac{16n}{n-4} + 32 \left(\frac{n}{n-4} \right)^3 \right] \left[\frac{32n}{n-4} \right]} / \theta. \quad (2.11)$$

To prove Proposition 2.3, we need the following lemma which estimates the second moment of S .

Lemma 2.4. *Let S be defined by (2.7) for some $m \in \{2, 3, 4\}$ and $n \geq 6$. Suppose $\gamma \leq 1/c_0$ where γ was defined in (1.4). Under the assumptions (1.1) and (2.6), we have*

$$\frac{n-1}{n-2} - \frac{2n}{(n-4)c_0^{2/3}} - \frac{24n}{(n-5)^2} \leq \mathbb{E}S^2 \leq \frac{n}{n-5} + \frac{24n}{(n-5)^2}. \quad (2.12)$$

Proof. Writing $\sum_{1 \leq i \neq j \leq n-m} \sum_{1 \leq k \neq l \leq n-m}$ as

$$\sum_{1 \leq i, j, k, l \leq n-m} - \sum_{1 \leq i=j \leq n-m} \sum_{1 \leq k, l \leq n-m} - \sum_{1 \leq i, j \leq n-m} \sum_{1 \leq k=l \leq n-m} + \sum_{1 \leq i=j \leq n-m} \sum_{1 \leq k=l \leq n-m}$$

and using the assumption (1.1),

$$\begin{aligned} \mathbb{E}S^2 &= \mathbb{E}\left(\sum_{i=1}^{n-m} X_{i\tau(i)}\right)^2 = \sum_{i=1}^{n-m} \mathbb{E}X_{i\tau(i)}^2 + \sum_{1 \leq i \neq j \leq n-m} \mathbb{E}X_{i\tau(i)}X_{j\tau(j)} \\ &= \frac{1}{n-m} \sum_{i,j=1}^{n-m} \mathbb{E}X_{ij}^2 + \frac{1}{(n-m)(n-m-1)} \sum_{1 \leq i \neq j \leq n-m} \sum_{1 \leq k \neq l \leq n-m} \mathbb{E}X_{ik}X_{jl} \\ &= \frac{1}{n-m} \sum_{i,j=1}^{n-m} (c_{ij}^2 + \sigma_{ij}^2) + \frac{1}{(n-m)(n-m-1)} \sum_{1 \leq i \neq j \leq n-m} \sum_{1 \leq k \neq l \leq n-m} c_{ik}c_{jl} \\ &= \frac{1}{n-m} \sum_{i,j=1}^{n-m} \sigma_{ij}^2 + \frac{1}{n-m-1} \sum_{i,j=1}^{n-m} c_{ij}^2 \\ &\quad + \frac{1}{(n-m)(n-m-1)} \sum_{i,j=1}^{n-m} c_{ij} \left(\sum_{k=n-m+1}^n c_{ik} + \sum_{l=n-m+1}^n c_{lj} + \sum_{k,l=n-m+1}^n c_{kl} \right). \end{aligned} \tag{2.13}$$

Under the assumption (2.6), $\mathbb{E}S^2$ is close to 1 intuitively. We quantify it as follows. From (1.1) and (3.1),

$$\left| \sum_{i,j=1}^{n-m} c_{ij} \left(\sum_{k=n-m+1}^n c_{ik} \right) \right| = \left| \sum_{i=1}^{n-m} \left(\sum_{k=n-m+1}^n c_{ik} \right)^2 \right| \leq \left| \sum_{i=1}^{n-m} m \sum_{k=n-m+1}^n c_{ik}^2 \right| \leq m(n-1).$$

Similarly,

$$\left| \sum_{i,j=1}^{n-m} c_{ij} \sum_{l=n-m+1}^n c_{lj} \right| \leq m(n-1).$$

Moreover,

$$\left| \sum_{i,j=1}^{n-m} c_{ij} \sum_{k,l=n-m+1}^n c_{kl} \right| = \left| \left(\sum_{k,l=n-m+1}^n c_{kl} \right)^2 \right| \leq m^2 \sum_{k,l=n-m+1}^n c_{kl}^2 \leq m^2(n-1).$$

Therefore, from (2.13),

$$\mathbb{E}S^2 \leq \frac{1}{n-m-1} \sum_{i,j=1}^n (\sigma_{ij}^2 + c_{ij}^2) + \frac{(2m+m^2)(n-1)}{(n-m)(n-m-1)} \leq \frac{n}{n-5} + \frac{24n}{(n-5)^2}.$$

Since $\gamma \leq 1/c_0$, using Hölder's inequality, we have

$$\sum_{\substack{1 \leq i, j \leq n: i > n-m \\ \text{or } j > n-m}} (\sigma_{ij}^2 + c_{ij}^2) = \sum_{\substack{1 \leq i, j \leq n: i > n-m \\ \text{or } j > n-m}} \mathbb{E} X_{ij}^2 \leq (8n)^{1/3} \left(\sum_{i,j=1}^n \mathbb{E} |X_{ij}|^3 \right)^{2/3} \leq 2n/c_0^{2/3}.$$

A lower bound can be obtained as

$$\begin{aligned} \mathbb{E} S^2 &\geq \frac{1}{n-m} \sum_{i,j=1}^{n-m} (\sigma_{ij}^2 + c_{ij}^2) - \frac{24n}{(n-5)^2} \\ &= \frac{1}{n-m} \sum_{i,j=1}^n (\sigma_{ij}^2 + c_{ij}^2) - \frac{1}{n-m} \sum_{\substack{1 \leq i, j \leq n: i > n-m \\ \text{or } j > n-m}} (\sigma_{ij}^2 + c_{ij}^2) - \frac{24n}{(n-5)^2} \\ &\geq \frac{n-1}{n-2} - \frac{2n}{(n-4)c_0^{2/3}} - \frac{24n}{(n-5)^2}. \end{aligned}$$

□

Proof of Proposition 2.3. For any $m \in \{2, 3, 4\}$, we construct an exchangeable pair (S, S') by uniformly selecting two different indices $I, J \in [n-m]$ and letting $S' = S - X_{I\tau(I)} - X_{J\tau(J)} + X_{I\tau(J)} + X_{J\tau(I)}$. An approximate linearity condition with an error term can be established as

$$\begin{aligned} &\mathbb{E}(S' - S|S) \\ &= \frac{1}{(n-m)(n-m-1)} \sum_{1 \leq i, j \leq n-m} \mathbb{E} \left\{ \left[X_{i\tau(j)} + X_{j\tau(i)} \right] - \left[X_{i\tau(i)} + X_{j\tau(j)} \right] \middle| S \right\} \\ &= \frac{1}{(n-m)(n-m-1)} \mathbb{E} \left\{ 2 \sum_{i,j=1}^{n-m} X_{ij} - 2(n-m)S \middle| S \right\} \\ &= -\lambda S + R \end{aligned} \tag{2.14}$$

where $\lambda = 2/(n-m-1)$ and

$$R = \frac{2}{(n-m)(n-m-1)} \mathbb{E} \left(\sum_{i,j=1}^{n-m} X_{ij} \middle| S \right).$$

To apply the concentration inequality in Lemma 2.1, we need to:

1. Bound $\sqrt{\mathbb{E} R^2}/\lambda$.
2. Bound

$$\delta = \frac{\mathbb{E} |S' - S|^3}{\lambda}. \tag{2.15}$$

3. Bound

$$\sqrt{\text{Var}\left(\mathbb{E}\left(\frac{1}{2\lambda}(S' - S)^2 I(|S' - S| \leq \delta) | S\right)\right)}.$$

Firstly,

$$\begin{aligned} \sqrt{\mathbb{E}R^2} &= \frac{2}{(n-m)(n-m-1)} \sqrt{\mathbb{E}(\mathbb{E}(\sum_{i,j=1}^{n-m} X_{ij} | S))^2} \\ &= \frac{2}{(n-m)(n-m-1)} \sqrt{\text{Var}(\mathbb{E}(\sum_{i,j=1}^{n-m} X_{ij} | S)) + (\mathbb{E} \sum_{i,j=1}^{n-m} X_{ij})^2} \\ &\leq \frac{2}{(n-m)(n-m-1)} \sqrt{\text{Var}(\sum_{i,j=1}^{n-m} X_{ij}) + (\sum_{i,j=1}^{n-m} c_{ij})^2} \\ &= \frac{2}{(n-m)(n-m-1)} \sqrt{\sum_{i,j=1}^{n-m} \sigma_{ij}^2 + (\sum_{i,j=n-m+1}^n c_{ij})^2} \\ &\leq \frac{2}{(n-m)(n-m-1)} \sqrt{\sum_{i,j=1}^{n-m} \sigma_{ij}^2 + m \sum_{i,j=n-m+1}^n c_{ij}^2} \\ &\leq \frac{2}{(n-m)(n-m-1)} \sqrt{m(\sum_{i,j=1}^n \sigma_{ij}^2 + \sum_{i,j=1}^n c_{ij}^2)} \\ &\leq \frac{2\sqrt{mn}}{(n-m)(n-m+1)} \end{aligned}$$

where we used the assumptions (1.1) and (2.6). Therefore,

$$\frac{\sqrt{\mathbb{E}R^2}}{\lambda} \leq \frac{2\sqrt{n}}{n-4}. \quad (2.16)$$

Next we bound δ of (2.15).

$$\begin{aligned} &\mathbb{E}|S' - S|^3 \\ &= \mathbb{E} \frac{1}{(n-m)(n-m-1)} \sum_{1 \leq i,j \leq n-m} \mathbb{E}(|X_{i\tau(i)} + X_{j\tau(j)} - X_{i\tau(j)} - X_{j\tau(i)}|^3 | \mathbb{X}) \\ &\leq \mathbb{E} \frac{16}{(n-m)(n-m-1)} \sum_{1 \leq i,j \leq n-m} \mathbb{E}(|X_{i\tau(i)}|^3 + |X_{j\tau(j)}|^3 + |X_{i\tau(j)}|^3 + |X_{j\tau(i)}|^3 | \mathbb{X}) \\ &\leq \frac{64n}{(n-m)(n-m-1)} \gamma \end{aligned}$$

where we used the fact that

$$\begin{aligned} & |X_{i\tau(j)} + X_{j\tau(i)} - X_{i\tau(i)} - X_{j\tau(j)}|^3 \\ & \leq 16(|X_{i\tau(j)}|^3 + |X_{j\tau(i)}|^3 + |X_{i\tau(i)}|^3 + |X_{j\tau(j)}|^3). \end{aligned} \quad (2.17)$$

Therefore,

$$\delta \leq \frac{32n}{n-4}\gamma. \quad (2.18)$$

Now we turn to the final step of bounding $\sqrt{\text{Var}\left(\mathbb{E}\left(\frac{1}{2\lambda}(S' - S)^2 I(|S' - S| \leq \delta) | S\right)\right)}$. Denote

$$\alpha_{ij}^\tau = (X_{i\tau(i)} + X_{j\tau(j)} - X_{i\tau(j)} - X_{j\tau(i)})^2 I(|X_{i\tau(i)} + X_{j\tau(j)} - X_{i\tau(j)} - X_{j\tau(i)}| \leq \delta).$$

We have

$$\frac{1}{2\lambda} \mathbb{E}\left((S' - S)^2 I(|S' - S| \leq \delta) | \mathbb{X}, \tau\right) = \frac{1}{4(n-m)} \sum_{1 \leq i \neq j \leq n-m} \alpha_{ij}^\tau.$$

Therefore, with $|\cdot|$ meaning the cardinality of a set,

$$\begin{aligned} & \text{Var}\left(\mathbb{E}\left(\frac{1}{2\lambda}(S' - S)^2 I(|S' - S| \leq \delta) | S\right)\right) \\ & \leq \text{Var}\left(\frac{1}{2\lambda} \mathbb{E}\left((S' - S)^2 I(|S' - S| \leq \delta) | \mathbb{X}, \tau\right)\right) \\ & = \frac{1}{16(n-m)^2} \left\{ 2 \sum_{1 \leq i \neq j \leq n-m} \text{Var}(\alpha_{ij}^\tau) \right. \\ & \quad + \sum_{1 \leq i, j, i', j' \leq n-m, i \neq j, i' \neq j', |i, j, i', j'| = 3} \text{Cov}(\alpha_{ij}^\tau, \alpha_{i'j'}^\tau) \\ & \quad \left. + \sum_{1 \leq i, j, i', j' \leq n-m, |i, j, i', j'| = 4} \text{Cov}(\alpha_{ij}^\tau, \alpha_{i'j'}^\tau) \right\} \\ & := R_1 + R_2 + R_3. \end{aligned}$$

The terms R_1 and R_2 are easy to bound.

$$|R_1| \leq \frac{2}{16(n-m)^2} \sum_{i, j=1}^{n-m} \delta \mathbb{E}|X_{i\tau(i)} + X_{j\tau(j)} - X_{i\tau(j)} - X_{j\tau(i)}|^3 \leq \frac{8n\delta}{(n-4)^2} \gamma. \quad (2.19)$$

From $\text{Cov}(X, Y) \leq (\text{Var}(X) + \text{Var}(Y))/2$, (2.17) and the restriction that $i \neq j, i' \neq j', |i, j, i', j'| =$

3, we have

$$\begin{aligned}
& |R_2| \\
& \leq \frac{\delta}{16(n-m)^2} \sum_{1 \leq i, j, i', j' \leq n-m, i \neq j, i' \neq j', |i, j, i', j'|=3} \mathbb{E} |X_{i\tau(i)} + X_{j\tau(j)} - X_{i\tau(j)} - X_{j\tau(i)}|^3 \quad (2.20) \\
& \leq \frac{16n\delta}{n-4} \gamma
\end{aligned}$$

where one factor of 64 comes from (2.17) by separating the four summands and the other factor 4 comes from the constraint $|i, j, i', j'| = 3$. Let $\alpha_{ij}^{kl} = (X_{ik} + X_{jl} - X_{il} - X_{jk})^2 I(|X_{ik} + X_{jl} - X_{il} - X_{jk}| \leq \delta)$. For $|i, j, i', j'| = 4$,

$$\begin{aligned}
& \text{Cov}(\alpha_{ij}^\tau, \alpha_{i'j'}^\tau) = \mathbb{E} \alpha_{ij}^\tau \alpha_{i'j'}^\tau - \mathbb{E} \alpha_{ij}^\tau \mathbb{E} \alpha_{i'j'}^\tau \\
& = \frac{1}{(n-m)(n-m-1)(n-m-2)(n-m-3)} \sum_{1 \leq k, l, k', l' \leq n-m, |k, l, k', l'|=4} \mathbb{E} \alpha_{ij}^{kl} \alpha_{i'j'}^{k'l'} \\
& \quad - \left[\frac{1}{(n-m)(n-m-1)} \sum_{1 \leq k \neq l \leq n-m} \mathbb{E} \alpha_{ij}^{kl} \right] \left[\frac{1}{(n-m)(n-m-1)} \sum_{1 \leq k' \neq l' \leq n-m} \mathbb{E} \alpha_{i'j'}^{k'l'} \right] \\
& = \frac{1}{(n-m)(n-m-1)(n-m-2)(n-m-3)} \sum_{1 \leq k, l, k', l' \leq n-m, |k, l, k', l'|=4} \mathbb{E} \alpha_{ij}^{kl} \mathbb{E} \alpha_{i'j'}^{k'l'} \\
& \quad - \frac{1}{(n-m)^2(n-m-1)^2} \sum_{1 \leq k, l, k', l' \leq n-m, |k, l, k', l'|=4} \mathbb{E} \alpha_{ij}^{kl} \mathbb{E} \alpha_{i'j'}^{k'l'} \\
& \quad - \frac{1}{(n-m)^2(n-m-1)^2} \sum_{1 \leq k, l, k', l' \leq n-m, k \neq l, k' \neq l', |k, l, k', l'| \leq 3} \mathbb{E} \alpha_{ij}^{kl} \mathbb{E} \alpha_{i'j'}^{k'l'} \\
& = \frac{4(n-m)-6}{(n-m)^2(n-m-1)^2(n-m-2)(n-m-3)} \sum_{1 \leq k, l, k', l' \leq n-m, |k, l, k', l'|=4} \mathbb{E} \alpha_{ij}^{kl} \mathbb{E} \alpha_{i'j'}^{k'l'} \\
& \quad - \frac{1}{(n-m)^2(n-m-1)^2} \sum_{1 \leq k, l, k', l' \leq n-m, k \neq l, k' \neq l', |k, l, k', l'| \leq 3} \mathbb{E} \alpha_{ij}^{kl} \mathbb{E} \alpha_{i'j'}^{k'l'}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
|R_3| &\leq \frac{1}{16(n-m)^2} \sum_{1 \leq i, j, i', j' \leq n-m, |i, j, i', j'|=4} \left[\frac{4(n-m)-6}{(n-m)^2(n-m-1)^2(n-m-2)(n-m-3)} \right. \\
&\quad \times \sum_{1 \leq k, l, k', l' \leq n-m, |k, l, k', l'|=4} \frac{\mathbb{E}(\alpha_{ij}^{kl})^2 + \mathbb{E}(\alpha_{i'j'}^{k'l'})^2}{2} \\
&\quad \left. + \frac{1}{(n-m)^2(n-m-1)^2} \sum_{1 \leq k, l, k', l' \leq n-m, k \neq l, k' \neq l', |k, l, k', l'| \leq 3} \frac{\mathbb{E}(\alpha_{ij}^{kl})^2 + \mathbb{E}(\alpha_{i'j'}^{k'l'})^2}{2} \right] \\
&\leq \frac{1}{16(n-m)^2} \sum_{1 \leq i, j, i', j' \leq n-m, |i, j, i', j'|=4} \left[\frac{4}{(n-m)(n-m-1)^2} \left(\sum_{1 \leq k \neq l \leq n-m} \frac{\mathbb{E}(\alpha_{ij}^{kl})^2}{2} \right. \right. \\
&\quad + \sum_{1 \leq k' \neq l' \leq n-m} \frac{\mathbb{E}(\alpha_{i'j'}^{k'l'})^2}{2} \left. \right) + \frac{4}{(n-m)(n-m-1)^2} \left(\sum_{1 \leq k \neq l \leq n-m} \frac{\mathbb{E}(\alpha_{ij}^{kl})^2}{2} \right. \\
&\quad \left. \left. + \sum_{1 \leq k' \neq l' \leq n-m} \frac{\mathbb{E}(\alpha_{i'j'}^{k'l'})^2}{2} \right) \right] \\
&\leq \frac{1}{2(n-m)^3(n-m-1)^2} \sum_{1 \leq i, j, i', j' \leq n-m, |i, j, i', j'|=4} \sum_{1 \leq k \neq l \leq n-m} \mathbb{E}(\alpha_{ij}^{kl})^2 \\
&\leq \frac{\delta}{2(n-m)^3} \sum_{i, j, k, l=1}^n \mathbb{E}|X_{ik} + X_{jl} - X_{il} - X_{jk}|^3 \leq 32 \left(\frac{n}{n-4} \right)^3 \delta \gamma.
\end{aligned} \tag{2.21}$$

From (2.19), (2.20), (2.21), and then applying (2.18), we obtain

$$\begin{aligned}
&\sqrt{\text{Var} \left(\mathbb{E} \left(\frac{1}{2\lambda} (S' - S)^2 I(|S' - S| \leq \delta) \middle| S \right) \right)} \\
&\leq \sqrt{\frac{8n}{(n-4)^2} + \frac{16n}{n-4} + 32 \left(\frac{n}{n-4} \right)^3 \sqrt{\delta \gamma}} \\
&\leq \sqrt{\frac{8n}{(n-4)^2} + \frac{16n}{n-4} + 32 \left(\frac{n}{n-4} \right)^3 \sqrt{\frac{32n}{n-4} \gamma}}.
\end{aligned} \tag{2.22}$$

Now we are ready to obtain a concentration inequality for S using Lemma 2.1. From (2.2),

and applying the bounds (2.16), (2.12), (2.18), (2.22), we obtain

$$\begin{aligned}
P(S \in [a, b]) &\leq \frac{\sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2} + \frac{2\sqrt{n}}{n-4}}}{\frac{n-1}{n-2} - \frac{2n}{(n-4)c_0^{2/3}} - \frac{24n}{(n-5)^2} - \frac{2\sqrt{n}}{n-4} \sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2}} - \frac{1}{2}} \left(\frac{b-a}{2} + \delta \right) \\
&\quad + \frac{\sqrt{\frac{8n}{(n-4)^2} + \frac{16n}{n-4} + 32\left(\frac{n}{n-4}\right)^3 \sqrt{\frac{32n}{n-4}}}}{\frac{n-1}{n-2} - \frac{2n}{(n-4)c_0^{2/3}} - \frac{24n}{(n-5)^2} - \frac{2\sqrt{n}}{n-4} \sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2}} - \frac{1}{2}} \gamma \\
&\leq c_1(b-a) + c_2\gamma.
\end{aligned}$$

□

Remark 2.5. From Proposition 2.3, the error in [16] can be corrected by conditioning on (using their notation)

$$\begin{aligned}
&J, K, L, M, \tau(J), \tau(K), \tau(L), \tau(M), \\
&\{\hat{X}_{ij} : i \in \{J, K, \tau^{-1}(L), \tau^{-1}(M)\}, j \in \{L, M, \tau(J), \tau(K)\}\},
\end{aligned}$$

and by applying our Proposition 2.3 instead of their Proposition 2.7.

3. Proof of the main result

From (1.1), $\mathbb{E}W = 0$. The variance of W can be calculated as follows. From (1.1),

$$\begin{aligned}
\text{Var}(W) &= \text{Var}\left(\sum_{i=1}^n X_{i\pi(i)}\right) \\
&= \sum_{i=1}^n \text{Var}(X_{i\pi(i)}) + \sum_{1 \leq i \neq j \leq n} \text{Cov}(X_{i\pi(i)}, X_{j\pi(j)}) \\
&= \sum_{i=1}^n \mathbb{E}(X_{i\pi(i)} - c_{i\cdot})^2 + \sum_{1 \leq i \neq j \leq n} \mathbb{E}(X_{i\pi(i)} - c_{i\cdot})(X_{j\pi(j)} - c_{j\cdot}) \\
&= \frac{1}{n} \sum_{i,j=1}^n \mathbb{E}(X_{ij} - c_{i\cdot})^2 + \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \sum_{1 \leq k \neq l \leq n} \mathbb{E}(X_{ik} - c_{i\cdot})(X_{jl} - c_{j\cdot}) \\
&= \frac{1}{n} \sum_{i,j=1}^n (\sigma_{ij}^2 + c_{ij}^2) + \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \sum_{1 \leq k \neq l \leq n} c_{ik}c_{jl} \\
&= \frac{1}{n} \sum_{i,j=1}^n (\sigma_{ij}^2 + c_{ij}^2) + \frac{1}{n(n-1)} \sum_{i,j=1}^n c_{ij}^2 \\
&= \frac{1}{n} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{n-1} \sum_{i,j=1}^n c_{ij}^2.
\end{aligned}$$

This proves the first part of the theorem. In the following, we work under the assumption that $\text{Var}(W) = 1$, i.e.,

$$\frac{1}{n} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{n-1} \sum_{i,j=1}^n c_{ij}^2 = 1 \quad (3.1)$$

We assume $\gamma \leq 1/447$, i.e., $c_0 = 447$ in Proposition 2.3. Otherwise the bound (1.3) is obviously true. From (3.1) and Hölder's inequality, we have

$$n-1 \leq \sum_{i,j=1}^n \mathbb{E}X_{ij}^2 \leq n^{2/3} \left(\sum_{i,j=1}^n \mathbb{E}|X_{ij}|^3 \right)^{2/3} = n^{4/3} \gamma^{2/3}. \quad (3.2)$$

Therefore, we have $n \geq 199800$ and (2.8) is satisfied. We prove Theorem 1.1 by applying the concentration inequality (2.9) in Stein's method. We follow the notation in Section 1 and construct an exchangeable pair (W, W') by uniformly selecting two different indices $I, J \in [n]$ (the ranges of I and J are different from those in the proof of Proposition 2.3) and let $W' = W - X_{I\pi(I)} - X_{J\pi(J)} + X_{I\pi(J)} + X_{J\pi(I)}$. Following the argument as in (2.14), we have

$$\mathbb{E}(W' - W|W) = -\lambda W + R \quad (3.3)$$

where $\lambda = 2/(n-1)$ and

$$R = \frac{2}{n(n-1)} \mathbb{E} \left(\sum_{i,j=1}^n X_{ij} | W \right).$$

The following bound on $\sqrt{\mathbb{E}R^2}$

$$\sqrt{\mathbb{E}R^2} \leq \frac{2}{n(n-1)} \sqrt{\text{Var} \left(\sum_{i,j=1}^n X_{ij} \right)} = \frac{2}{n(n-1)} \sqrt{\sum_{i,j=1}^n \sigma_{ij}^2} \leq \frac{2}{(n-1)\sqrt{n}} \quad (3.4)$$

is obtained by using the assumptions (1.1) and (3.1).

From the fact that (W, W') is an exchangeable pair and satisfies an approximate linearity condition (3.3), the following functional identity can be proved following the same argument as in (2.5).

$$\mathbb{E}Wf(W) = \frac{1}{2\lambda} \mathbb{E}(W' - W)(f(W') - f(W)) + \frac{\mathbb{E}Rf(W)}{\lambda}. \quad (3.5)$$

Let f be the bounded solution to the Stein equation

$$f'(w) - wf(w) = I(w \leq z) - \Phi(z). \quad (3.6)$$

It is known that (Chen and Shao [9])

$$|f(w)| \leq \frac{\sqrt{2\pi}}{4}, \quad |f'(w)| \leq 1 \quad \forall w \in \mathbb{R} \quad (3.7)$$

and

$$|(w+u)f(w+u) - (w+v)f(w+v)| \leq (|w| + \frac{\sqrt{2\pi}}{4})(|u| + |v|). \quad (3.8)$$

From (3.6) and (3.5), what we need to bound is

$$\begin{aligned} & \mathbb{P}(W \leq z) - \Phi(z) \\ &= \mathbb{E}f'(W) - \mathbb{E}Wf(W) \\ &= \mathbb{E}f'(W) \left(1 - \frac{(W' - W)^2}{2\lambda}\right) + \frac{1}{2\lambda} \mathbb{E}(W' - W) \int_0^{W' - W} (f'(W) - f'(W+t)) dt \\ &\quad - \frac{\mathbb{E}Rf(W)}{\lambda} \\ &:= R_1 + R_2 - R_3 \end{aligned}$$

From (3.4) and (3.7), and recalling $\lambda = 2/(n-1)$, we have

$$|R_3| \leq 1/\sqrt{n}. \quad (3.9)$$

To bound R_1 and R_2 , we need the concentration inequality obtained in the last section.

From (3.3) and (3.4),

$$\mathbb{E}(W' - W)^2 = 2\lambda - 2\mathbb{E}WR \begin{cases} \leq 2\lambda + 2\sqrt{\mathbb{E}R^2} \leq \frac{4}{n-1}(1 + \frac{1}{\sqrt{n}}). \\ \geq 2\lambda - 2\sqrt{\mathbb{E}R^2} \geq \frac{4}{n-1}(1 - \frac{1}{\sqrt{n}}). \end{cases} \quad (3.10)$$

We bound R_2 first. From (3.6),

$$\begin{aligned} R_2 &= \frac{1}{2\lambda} \mathbb{E}(W' - W) \int_0^{W' - W} (Wf(W) - (W+t)f(W+t))dt \\ &\quad + \frac{1}{4n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}(X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}) \\ &\quad \times \int_0^{X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}} \left[I(U \leq z - X_{i\pi(i)} - X_{j\pi(j)}) \right. \\ &\quad \left. - I(U \leq z - X_{i\pi(i)} - X_{j\pi(j)} - t) \right] dt \\ &:= R_{2,1} + R_{2,2} \end{aligned}$$

where $U = \sum_{1 \leq k \leq n: k \neq i, j} X_{k\pi(k)}$. Noting that U is independent of $\{X_{i\pi(i)}, X_{j\pi(j)}, X_{i\pi(j)}, X_{j\pi(i)}\}$ given $\pi(i), \pi(j)$, and that the conditional distribution of U given $\pi(i), \pi(j)$ is the same as the distribution of S in (2.7) (except for the indices but this does not matter), we can apply the concentration inequality (2.9) to obtain the following upper bound on $|R_{2,2}|$.

$$\begin{aligned} |R_{2,2}| &\leq \frac{1}{4n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}(X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}) \\ &\quad \times \int_0^{X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}} I(z - X_{i\pi(i)} - X_{j\pi(j)} - (t \vee 0) \leq U \leq z - X_{i\pi(i)} - X_{j\pi(j)} - (t \wedge 0)) dt \\ &\leq \frac{1}{4n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}(X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}) \\ &\quad \times \int_0^{X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}} (c_1|t| + c_2\gamma) dt \\ &= \frac{c_1}{8n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}|X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}|^3 \\ &\quad + \frac{c_2\gamma}{4n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}|X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}|^2 \\ &\leq 8c_1\gamma + c_2\gamma(1 + \frac{1}{\sqrt{n}}). \end{aligned}$$

In the last inequality, we used (2.17) and

$$\begin{aligned} \sum_{1 \leq i \neq j \leq n} \mathbb{E}|X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}|^2 &= n(n-1)\mathbb{E}(W' - W)^2 \\ &\leq 4n(1 + \frac{1}{\sqrt{n}}) \end{aligned}$$

where (3.10) is used in the last inequality. For $R_{2,1}$, from the property (3.8) of f with $w = U, u = X_{i\pi(i)} + X_{j\pi(j)}, v = X_{i\pi(i)} + X_{j\pi(j)} + t$,

$$\begin{aligned} |R_{2,1}| &\leq \frac{1}{4n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}(X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}) \\ &\quad \times \int_0^{X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}} (|U| + \frac{\sqrt{2\pi}}{4})(2|X_{i\pi(i)} + X_{j\pi(j)}| + t) dt \\ &= \frac{1}{4n} \sum_{1 \leq i \neq j \leq n} \mathbb{E}(|U| + \frac{\sqrt{2\pi}}{4}) \left[(X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)})^2 2|X_{i\pi(i)} + X_{j\pi(j)}| \right. \\ &\quad \left. + \frac{|X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}|^3}{2} \right] \\ &\leq 24(\sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2}} + \frac{\sqrt{2\pi}}{4})\gamma \end{aligned}$$

where we used (2.12), (2.17) and

$$\begin{aligned} &|X_{i\pi(j)} + X_{j\pi(i)} - X_{i\pi(i)} - X_{j\pi(j)}|^2 |X_{i\pi(i)} + X_{j\pi(j)}| \\ &\leq \frac{16}{3}(|X_{i\pi(j)}|^3 + |X_{j\pi(i)}|^3) + \frac{32}{3}(|X_{i\pi(i)}|^3 + |X_{j\pi(j)}|^3). \end{aligned}$$

Therefore, with

$$\begin{aligned} c_3 &:= \sqrt{\frac{n}{n-5} + \frac{24n}{(n-5)^2}}, \\ |R_2| &\leq (8c_1 + c_2(1 + \frac{1}{\sqrt{n}}) + 24c_3 + 6\sqrt{2\pi})\gamma. \end{aligned} \tag{3.11}$$

Next, we bound R_1 .

$$\begin{aligned}
R_1 &= \mathbb{E} f'(W) \left(1 - \frac{(W' - W)^2}{2\lambda}\right) \\
&= \frac{1}{n^2(n-1)^2} \\
&\quad \times \sum_{1 \leq i, j, k, l \leq n, i \neq j, k \neq l} \mathbb{E} \left(f'(W) \left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right) \middle| I = i, J = j, \pi(i) = k, \pi(j) = l \right) \\
&= \frac{1}{n^2(n-1)^2} \sum_{1 \leq i, j, k, l \leq n, i \neq j, k \neq l} \mathbb{E} \left(f'(W) \left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right) \middle| \pi(i) = k, \pi(j) = l \right)
\end{aligned}$$

since \mathbb{X}, π and (I, J) are independent. For each choice of $i \neq j, k \neq l$, let $\mathbb{X}^{ijkl} := \{X_{i'j'}^{ijkl} : i', j' \in [n]\}$ be the same as \mathbb{X} except that $\{X_{ik}, X_{il}, X_{jk}, X_{jl}\}$ has been replaced by an independent copy $\{X'_{ik}, X'_{il}, X'_{jk}, X'_{jl}\}$. Define

$$W^{ijkl} = \sum_{i'=1}^n X_{i'\pi(i')}^{ijkl}.$$

Then,

$$W^{ijkl} \text{ is independent of } \{X_{ik}, X_{il}, X_{jk}, X_{jl}\} \text{ and } \mathcal{L}(W^{ijkl}) = \mathcal{L}(W). \quad (3.12)$$

Next, we define a new permutation π_{ijkl} coupled with π such that

$$\mathcal{L}(\pi_{ijkl}) = \mathcal{L}(\pi | \pi(i) = k, \pi(j) = l).$$

This coupling has been constructed by Chen [3] and also by Goldstein [11]. We describe this coupling as follows.

$$\pi_{ijkl}(\alpha) = \begin{cases} \pi(\alpha), & \alpha \neq i, j, \pi^{-1}(k), \pi^{-1}(l) \\ k, & \alpha = i \\ l, & \alpha = j \\ \pi(i), & \alpha = \pi^{-1}(k) \\ \pi(j), & \alpha = \pi^{-1}(l). \end{cases}$$

Let

$$W_{ijkl} = \sum_{i'=1}^n X_{i'\pi_{ijkl}(i')}.$$

Since W_{ijkl} has the same distribution as the conditional distribution of W given $\pi(i) = k, \pi(j) = l$, and since \mathbb{X} and π are independent, we have

$$\begin{aligned} & \mathbb{E}\left(f'(W)\left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right) \middle| \pi(i) = k, \pi(j) = l\right) \\ &= \mathbb{E}f'(W_{ijkl})\left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} R_1 &= \mathbb{E}f'(W)\left(1 - \frac{(W' - W)^2}{2\lambda}\right) \\ &= \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E}\left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right) (f'(W_{ijkl}) - f'(W^{ijkl})) \\ &\quad + \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E}\left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right) f'(W^{ijkl}). \end{aligned}$$

Define index sets $\mathcal{I} = \{i, j, \pi^{-1}(k), \pi^{-1}(l)\}$ and $\mathcal{J} = \{k, l, \pi(i), \pi(j)\}$. Then $|\mathcal{I}| = |\mathcal{J}| \in \{2, 3, 4\}$. Letting $S = \sum_{i' \notin \mathcal{I}} X_{i'\pi(i')}$, we can write

$$W_{ijkl} = S + \sum_{i' \in \mathcal{I}} X_{i'\pi_{ijkl}(i')}$$

and

$$W^{ijkl} = S + \sum_{i' \in \mathcal{I}} X_{i'\pi(i')}^{ijkl}.$$

Since S is a function depending only on the components of \mathbb{X} outside the square $\mathcal{I} \times \mathcal{J}$,

$$\begin{aligned} & S \text{ is independent of } \{X_{il}, X_{jk}, X_{ik}, X_{jl}, \sum_{i' \in \mathcal{I}} X_{i'\pi_{ijkl}(i')}, \sum_{i' \in \mathcal{I}} X_{i'\pi(i')}^{ijkl}\} \\ & \text{given } \pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j). \end{aligned} \tag{3.13}$$

The conditional distribution of S given $\pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j)$ is the same as the distribution of S in (2.7) (except for the indices but this does not matter). From (2.12), $\mathbb{E}(|S| | \pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j)) \leq c_3$. From (3.12), (3.7) and (3.10),

$$\begin{aligned} & \left| \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E}\left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda}\right) f'(W^{ijkl}) \right| \\ &= \frac{1}{(n(n-1))^2} \left| \mathbb{E}f'(W) \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E}\left(1 - \frac{X_{il} + X_{jk} - X_{ik} - X_{jl}}{2\lambda}\right) \right| \\ &= \left| \mathbb{E}f'(W) \mathbb{E}\left(1 - \frac{(W' - W)^2}{2\lambda}\right) \right| \leq \frac{1}{\sqrt{n}}. \end{aligned}$$

Therefore,

$$\begin{aligned}
& |R_1| \\
& \leq \left| \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right) (f'(W_{ijkl}) - f'(W^{ijkl})) \right| \\
& \quad + \frac{1}{\sqrt{n}} \\
& = \left| \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right) \right. \\
& \quad \times \left. \left(f'(S + \sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}) - f'(S + \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}) \right) \right| + \frac{1}{\sqrt{n}} \\
& \leq R_{1,1} + R_{1,2} + \frac{1}{\sqrt{n}}
\end{aligned}$$

where using the Stein equation (3.6),

$$\begin{aligned}
R_{1,1} = & \left| \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right) \right. \\
& \times \left. \left((S + \sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}) f(S + \sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}) - (S + \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}) f(S + \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}) \right) \right|,
\end{aligned}$$

and

$$\begin{aligned}
R_{1,2} = & \left| \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right) \right. \\
& \times \left. \left(I(S + \sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}) \leq z \right) - I(S + \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}) \leq z \right) \Big|.
\end{aligned}$$

Applying (3.8), (3.13), (2.12) and (3.1), $R_{1,1}$ can be bounded as follows.

$$\begin{aligned}
R_{1,1} & \leq \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E}(\mathbb{E}(|S| | \pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j)) + \frac{\sqrt{2\pi}}{4}) \\
& \quad \times \mathbb{E} \left\{ \left(\left| \sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')} \right| + \left| \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl} \right| \right) \left| 1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right| \middle| \pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j) \right\} \\
& \leq \max \left\{ \frac{8(c_3 + \sqrt{2\pi}/4)}{n^2} \sum_{i,k=1}^n \mathbb{E}|X_{ik}|, 32(c_3 + \frac{\sqrt{2\pi}}{4}) \frac{n-1}{n} \gamma \right\} \\
& \leq \max \left\{ 8(c_3 + \frac{\sqrt{2\pi}}{4}) \frac{1}{\sqrt{n}}, 32(c_3 + \frac{\sqrt{2\pi}}{4}) \frac{n-1}{n} \gamma \right\}
\end{aligned}$$

where we used

$$\left| 1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right| \leq \max \left\{ 1, \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right\}. \quad (3.14)$$

Now we bound $R_{1,2}$.

$$\begin{aligned}
R_{1,2} &= \left| \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left\{ \mathbb{E} \left[\left(1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right) \right. \right. \right. \\
&\quad \left. \left. \times \left(I(S + \sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')} \leq z) - I(S + \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl} \leq z) \right) \middle| \pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j) \right] \right\} \right| \\
&\leq \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left\{ \mathbb{E} \left[\left| 1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right| \right. \right. \\
&\quad \left. \left. \times I(z - \max\{\sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}, \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}\} \leq S \leq z - \min\{\sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}, \sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}\}) \right. \right. \\
&\quad \left. \left. \left| \pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j) \right] \right\}.
\end{aligned}$$

Recall (3.13) and the fact that the conditional distribution of S given $\pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j)$ is the same as the distribution of S in (2.7) (except for the indices but this does not matter). We can therefore apply the concentration inequality (2.9) to obtain the following upper bound on $R_{1,2}$.

$$\begin{aligned}
R_{1,2} &\leq \frac{1}{(n(n-1))^2} \sum_{1 \leq i,j,k,l \leq n, i \neq j, k \neq l} \mathbb{E} \left| 1 - \frac{(X_{il} + X_{jk} - X_{ik} - X_{jl})^2}{2\lambda} \right| \\
&\quad \times \{c_1(|\sum_{i' \in \mathcal{I}} X_{i' \pi_{ijkl}(i')}| + |\sum_{i' \in \mathcal{I}} X_{i' \pi(i')}^{ijkl}|) + c_2\gamma\} \\
&\leq c_1 \max\left\{\frac{8}{\sqrt{n}}, 32\frac{n-1}{n}\gamma\right\} + c_2(1 + \frac{1}{\sqrt{n}})\gamma
\end{aligned}$$

where we used (3.14) and (3.10). Therefore,

$$|R_1| \leq (c_1 + c_3 + \frac{\sqrt{2\pi}}{4}) \max\left\{\frac{8}{\sqrt{n}}, 32\frac{n-1}{n}\gamma\right\} + c_2(1 + \frac{1}{\sqrt{n}})\gamma + \frac{1}{\sqrt{n}}. \quad (3.15)$$

Summing (3.15), (3.11), (3.9) yields an upper bound on $\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)|$ as

$$\begin{aligned}
&\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \\
&\leq \left(40c_1 + 2(1 + \frac{1}{\sqrt{n}})c_2 + 14\sqrt{2\pi} + 56c_3 + 2(\frac{n}{n-1})^{3/2} \right) \gamma
\end{aligned} \quad (3.16)$$

where we used (3.2). Recall $c_0 = 447$, $n \geq 199800$. Using $c_0 = 447$ and $n = 199800$, the upper bound in (3.16) is calculated to be smaller than 447γ . Since c_1, c_2 and c_3 decrease as n increases, (1.3) holds for $n \geq 199800$. For $n < 199800$, $\gamma > 1/447$ and (1.3) holds trivially. This completes the proof of Theorem 1.1.

Acknowledgements

This work is based on part of the Ph.D. thesis of the second author. The second author is thankful to the first author for his guidance and helpful discussions. Both the authors would like to thank an anonymous referee for helpful comments. This work is partially supported by Grant C-389-000-010-101 and Grant C-389-000-012-101 at the National University of Singapore.

References

- [1] BARBOUR, A.D. AND CHEN, L.H.Y. (2005) *An Introduction to Stein's method*. Lecture Notes Series **4**, Institute for Mathematical Sciences, National University of Singapore, Singapore University Press and World Scientific.
- [2] BOLTHAUSEN, E. (1984) An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **66** 379-386.
- [3] CHEN, L.H.Y. (1975) An approximation theorem for sums of certain randomly selected indicators. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **33** 69-74.
- [4] CHEN, L.H.Y. (1986) The rate of convergence in a central limit theorem for dependent random variables with arbitrary index set. *IMA Preprint Series* **243**, Univ. Minnesota.
- [5] CHEN, L.H.Y. (1998) Stein's method: some perspectives with applications. *Probability Towards 2000*. L. Accardi and C.C. Heyde, eds., Lecture Notes in Statistics **128**, Springer Verlag, 515-528.
- [6] CHEN, L.H.Y., GOLDSTEIN, L. AND SHAO, Q.M. (2011) *Normal approximation by Stein's method*. Springer.
- [7] CHEN, L.H.Y. AND SHAO, Q.M. (2001) A non-uniform Berry-Esseen bound via Stein's method. *Probab. Theory Related Fields* **120** 236-254.
- [8] CHEN, L.H.Y. AND SHAO, Q.M. (2004) Normal approximation under local dependence. *Ann. Probab.* **32** 1985-2028.
- [9] CHEN, L.H.Y. AND SHAO, Q.M. (2005). Stein's method for normal approximation. *An Introduction to Stein's Method*. A.D. Barbour and L. H. Y. Chen, eds., Lecture Notes Series **4**, Institute for Mathematical Sciences, National University of Singapore, Singapore University Press and World Scientific, 1-59.

- [10] DIACONIS, P. (1977) The distribution of leading digits and uniform distribution mod 1. *Ann. Probab.* **5** 72-81.
- [11] GOLDSTEIN, L. (2005) Berry-Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J. Appl. Probab.* **42** 661-683.
- [12] GOLDSTEIN, L. (2007) L^1 bounds in normal approximation. *Ann. Probab.* **35** 1888-1930.
- [13] HO, S.T. AND CHEN, L.H.Y. (1978) An L_p bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.* **6** 231-249.
- [14] HOEFFDING, W. (1951) A combinatorial central limit theorem. *Annals of Mathematical Statistics* **22** 558-566.
- [15] JOHNSON, W.B. AND LINDENSTRAUSS, J. (1984) Extensions of Lipschitz mappings into a Hilbert space. *Contemp. Math.* **26** 189-206.
- [16] NEAMMANEE, K. AND SUNTORNOCHOST, J. (2006) A uniform bound on a combinatorial central limit theorem. *Stoch. Anal. Appl.* **3** 559-578.
- [17] RINOTT, Y. AND ROTAR, V. (1997) On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U -statistics. *Ann. Appl. Probab.* **7** 1080-1105.
- [18] STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.* **2** Univ. California Press. Berkeley, Calif., 583-602.
- [19] STEIN, C. (1986). *Approximate Computation of Expectations*. Lecture Notes 7, Inst. Math. Statist., Hayward, Calif.
- [20] WALD, A. AND WOLFOWITZ, J. (1944) Statistical tests based on permutations of the observations. *Annals of Mathematical Statistics* **15** 358-372.
- [21] WOLFF, P. (2012) On randomness reduction in the Johnson-Lindenstrauss lemma. *Preprint*.